

Attachment  
"A"

SEQUENCE LISTING

<110> McIninch, James

<120> COMPUTATIONAL NUCLEIC ACID CODING AND FEATURE ANALYSIS

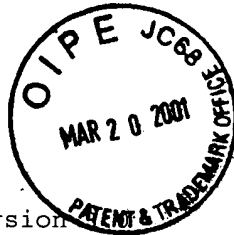
<130> 04983.0220.00US00

<140> 09/698,213

<141> 2000-10-30

<160> 13

<170> PatentIn version



<210> 1

<211> 2165

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> unsure

<222> (1)...(2165)

<223> Unsure at all n locations

<220>

<223> Ecotype Landsberg, genomic DNA

<400> 1

tactcaaaaa tatattccat gcttaattag gccggattcg cggtgacgat gcaccaagag	60
cggtttttcc gagcattgta ggccgtcctc gccacaccgg tgtgatgggt gggatgggac	120
aaaaggatgc ttatgttggg gacgaggctc aatcaaaacg tggatatcttg actctgaagt	180
accaattga gcatggaatt gttaataatt gggatgacat ggagaagatt tggcatcaca	240
ctttctacaa tgagcttcgt gttgccctg aagaacatcc ggttctcttg accgaagctc	300
ctctcaatcc gaaagctaac cgtgagaaga tgactcagat catgtttgag acattcaata	360
ctcctgctat gtatgttgcc attcaagctg ttctctcact ctatgccagt ggccgtacta	420
ctggtcagta cttactaca ttctttttat accgtttggg tgaaataaaa ttcggtttgg	480
ttcgattcga gtttgctctc attattttta ttttgttggg taggtattgt tttggactcc	540
ggagatggtg tgagccacac ggtaccaatc tacgagggtt atgcattcc acacgcaatc	600
ctgcgtcttg atcttgagg tcgtgaccta accgaccacc ttatgaaaat cctgacagag	660
cgtggttact ctttcaccac aactgctgag cgtgagattg ttagagacat gaaggagaag	720
ctctcttaca ttgccttggg ctttgaacaa gagctcgaga cttccaaaac aagctcatcc	780
gttgagaaga gcttcgagct gccagacggt caagtgatca ccatcggggc agagcgtttc	840

cgatgccctg aagttctgtt tcagccatcg atgatcggaa tggaaaatcc ggggaattcat	900
gaaactactt acaactcaat catgaaatgt gatgtggata tcaggaagga tctttatgga	960
aacattgtgc ttagtggtgg caccacaatg ttcgatggga ttggtgatag gatgagtaaa	1020
gagatcacag cgttggtccc aagcagtatg aacatcaaag tggtggtccc accggaaagg	1080
aagtacagtg tctggatcgg tggctctatc ttggcttccc tcagtacttt ccagcaggta	1140
aattacttac tataacttaat acataaagtc tattagtgat ttgatgtata aagtgttaca	1200
aaaatgtgtt ccaaatttgc agatgtggat tgcgaaagcg gagtatgatg aatctggacc	1260
gtcaatcgtc cacaggaagt gcttctgac aaaagtcacc aagtaaaaca agagcggtaa	1320
aaattttgat atcagttttt caccctgaag ccagttgcta taattactca caacttctct	1380
atttgtgttc ttttattctt gtccctcggt gttcatttta atctcttttt tgcaacaaag	1440
caacttaaaa aaacagagca gtcattaaca gaatgttatt attatatata tgtatacata	1500
ttagtataca ccattatctt cattaaaaca tttatcatat aaggatagga ttctatacat	1560
cgatatatct attttggtga cactattcag cacatgctta tgtcttatct tgttagtata	1620
tgtaaccaa gacaaataat agatgctaca aattgttttc tttgaagcaa aaatttcaat	1680
cttaaaattg tttttttcca gggtacacaa aaaaaacttg tagtttgtaa attttctata	1740
caattttggg gatctcaaca agaacatgaa cttcaacttc tagtcatatg acgacctgag	1800
tctgcgcggc tgtgaatctc tttgctgcag taaatgttta caagtgggtgt gtaaatgggt	1860
actgattcaa aagctttaag aaatctacac atttcgtgaa attatttagc agacttgata	1920
ttaaaaatct aggataaaat gactatccaa agacaaatag gactgtttca catgttcccc	1980
tgattcttgt agtcataac tcatcagcag ttaacttttc tacctcatac acgctcgcaa	2040
tncgtttgga attatcagct ntaatttttc taattctttg gaaattatta gcagctcgat	2100
caaatggggc atggcttctt cttctatctg caactcatct aaactttcca tgaagaaaca	2160
aagct	2165

<210> 2  
 <211> 6  
 <212> PRT  
 <213> Unknown

<220>  
 <223> Describes a predicted protein sequence

<400> 2  
 Arg Phe Phe Arg Ala Leu

1

5

<210> 3  
<211> 7  
<212> PRT  
<213> Unknown

<220>  
<223> Describes a predicted protein sequence

<400> 3  
Ala Val Leu Ala Thr Pro Val  
1 5

<210> 4  
<211> 21  
<212> PRT  
<213> Unknown

<220>  
<223> Describes a predicted protein sequence

<400> 4  
Trp Leu Gly Trp Asp Lys Arg Met Leu Met Leu Glu Thr Arg Leu Asn  
1 5 10 15

Gln Asn Val Val Ser  
20

<210> 5  
<211> 35  
<212> PRT  
<213> Unknown

<220>  
<223> Describes a predicted protein sequence

<400> 5  
Ser Thr Gln Leu Ser Met Glu Leu Leu Ile Ile Gly Met Thr Trp Arg  
1 5 10 15

Arg Phe Gly Ile Thr Leu Ser Thr Met Ser Phe Val Leu Pro Leu Lys  
20 25 30

Asn Ile Arg  
35

<210> 6  
<211> 56  
<212> PRT  
<213> Unknown

<220>

<223> Describes a predicted protein sequence

<400> 6

Leu Thr Glu Ala Pro Leu Asn Pro Lys Ala Asn Arg Glu Lys Met Thr  
1 5 10 15  
Gln Ile Met Phe Glu Thr Phe Asn Thr Pro Ala Met Tyr Val Ala Ile  
20 25 30  
Gln Ala Val Leu Ser Leu Tyr Ala Ser Gly Arg Thr Thr Gly Gln Tyr  
35 40 45  
Ile Thr Thr Phe Phe Leu Tyr Arg  
50 55

<210> 7

<211> 191

<212> PRT

<213> Unknown

<220>

<223> Describes a predicted protein sequence

<400> 7

Ser Gly Asp Gly Val Ser His Thr Val Pro Ile Tyr Glu Gly Tyr Ala  
1 5 10 15  
Leu Pro His Ala Ile Leu Arg Leu Asp Leu Ala Gly Arg Asp Leu Thr  
20 25 30  
Asp His Leu Met Lys Ile Leu Thr Glu Arg Gly Tyr Ser Phe Thr Thr  
35 40 45  
Thr Ala Glu Arg Glu Ile Val Arg Asp Met Lys Glu Lys Leu Ser Tyr  
50 55 60  
Ile Ala Leu Asp Phe Glu Gln Glu Leu Glu Thr Ser Lys Thr Ser Ser  
65 70 75 80  
Ser Val Glu Lys Ser Phe Glu Leu Pro Asp Gly Gln Val Ile Thr Ile  
85 90 95  
Gly Ala Glu Arg Phe Arg Cys Pro Glu Val Leu Phe Gln Pro Ser Met  
100 105 110  
Ile Gly Met Glu Asn Pro Gly Ile His Glu Thr Thr Tyr Asn Ser Ile  
115 120 125  
Met Lys Cys Asp Val Asp Ile Arg Lys Asp Leu Tyr Gly Asn Ile Val  
130 135 140  
Leu Ser Gly Gly Thr Thr Met Phe Asp Gly Ile Gly Asp Arg Met Ser  
145 150 155 160  
Lys Glu Ile Thr Ala Leu Ala Pro Ser Ser Met Lys Ile Lys Val Val  
165 170 175

Ala Pro Pro Glu Arg Lys Tyr Ser Val Trp Ile Gly Gly Ser Ile  
180 185 190

<210> 8  
<211> 13  
<212> PRT  
<213> Unknown

<220>  
<223> Describes a predicted protein sequence

<400> 8  
Val Pro Asn Leu Gln Met Trp Ile Ala Lys Ala Glu Tyr  
1 5 10

<210> 9  
<211> 26  
<212> PRT  
<213> Unknown

<220>  
<223> Describes a predicted protein sequence

<400> 9  
Asn Leu Asp Arg Gln Ser Ser Thr Gly Ser Ala Ser Asp Gln Lys Ser  
1 5 10 15

Pro Ser Lys Thr Arg Ala Val Lys Ile Leu  
20 25

<210> 10  
<211> 56  
<212> PRT  
<213> Unknown

<220>  
<223> Describes a predicted protein sequence

<400> 10

Asn Ser Ser Ala Val Asn Phe Ser Thr Ser Tyr Thr Leu Ala Ile Arg  
1 5 10 15

Leu Glu Leu Ser Ala Leu Ile Phe Leu Ile Ser Leu Glu Ile Ile Ser  
20 25 30

Ser Ser Ile Lys Trp Gly Met Ala Ser Ser Ser Ile Cys Asn Ser Ser  
35 40 45

Lys Leu Ser Met Lys Lys Gln Ser  
50 55

<210> 11  
<211> 194  
<212> PRT  
<213> Unknown

<220>

<223> Describes a predicted protein sequence

<400> 11

Ser Gly Asp Gly Val Ser His Thr Val Pro Ile Tyr Glu Gly Tyr Ala  
1 5 10 15

Leu Pro His Ala Ile Leu Arg Leu Asp Leu Ala Gly Arg Asp Leu Thr  
20 25 30

Asp His Leu Met Lys Ile Leu Thr Glu Arg Gly Tyr Ser Phe Thr Thr  
35 40 45

Thr Ala Glu Arg Glu Ile Val Arg Asp Met Lys Glu Lys Leu Ser Tyr  
50 55 60

Ile Ala Leu Asp Phe Glu Gln Glu Leu Glu Thr Ser Lys Thr Ser Ser  
65 70 75 80

Ser Val Glu Lys Ser Phe Glu Leu Pro Asp Gly Gln Val Ile Thr Ile  
85 90 95

Gly Ala Glu Arg Phe Arg Cys Pro Glu Val Leu Phe Gln Pro Ser Met  
100 105 110

Ile Gly Met Glu Asn Pro Gly Ile His Glu Thr Thr Tyr Asn Ser Ile  
115 120 125

Met Lys Cys Asp Val Asp Ile Arg Lys Asp Leu Tyr Gly Asn Ile Val  
130 135 140

Leu Ser Gly Gly Thr Thr Met Phe Asp Gly Ile Gly Asp Arg Met Ser  
145 150 155 160

Lys Glu Ile Thr Ala Leu Ala Pro Ser Ser Met Lys Ile Lys Val Val  
165 170 175

Ala Pro Pro Glu Arg Lys Tyr Ser Val Trp Ile Gly Gly Ser Ile Leu  
180 185 190

Ala Ser

<210> 12  
<211> 9  
<212> PRT  
<213> Unknown

<220>

<223> Describes a predicted protein sequence

<400> 12

Gln Met Trp Ile Ala Lys Ala Glu Tyr  
1 5

<210> 13  
<211> 296  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> Ecotype columbia, describes actin

<400> 13

Met Glu Lys Ile Trp His His Thr Phe Tyr Asn Glu Leu Arg Val Ala  
1 5 10 15

Pro Glu Glu His Pro Val Leu Leu Thr Glu Ala Pro Leu Asn Pro Lys  
20 25 30

Ala Asn Arg Glu Lys Met Thr Gln Ile Met Phe Glu Thr Phe Asn Thr  
35 40 45

Pro Ala Met Tyr Val Ala Ile Gln Ala Val Leu Ser Leu Ala Ser Gly  
50 55 60

Arg Thr Thr Gly Gly Ile Val Leu Asp Ser Gly Asp Gly Val Ser His  
65 70 75 80

Thr Val Pro Ile Tyr Glu Gly Tyr Ala Leu Pro His Ala Ile Leu Arg  
85 90 95

Leu Asp Leu Ala Gly Arg Asp Leu Thr Asp His Leu Met Lys Ile Leu  
100 105 110

Thr Glu Arg Gly Tyr Ser Phe Thr Thr Thr Ala Glu Arg Glu Ile Val  
115 120 125

Arg Asp Met Lys Glu Lys Leu Ser Tyr Ile Ala Leu Asp Phe Glu Gln  
130 135 140

Glu Leu Glu Thr Ser Lys Thr Ser Ser Ser Val Glu Lys Ser Phe Glu  
145 150 155 160

Leu Pro Asp Gly Gln Val Ile Thr Ile Gly Ala Glu Arg Phe Arg Cys  
165 170 175

Pro Glu Val Leu Phe Gln Pro Ser Met Ile Gly Met Glu Asn Pro Gly  
180 185 190

Ile His Glu Thr Thr Tyr Asn Ser Ile Met Lys Cys Asp Val Asp Ile  
195 200 205

Arg Lys Asp Leu Tyr Gly Asn Ile Val Leu Ser Gly Gly Thr Thr Met  
210 215 220

Phe Gly Gly Ile Gly Asp Arg Met Ser Lys Glu Ile Thr Ala Leu Ala

225		230		235		240
Pro Ser Ser Met Lys Ile Lys Val Val Ala Pro Pro Glu Arg Lys Tyr						
	245		250		255	
Ser Val Trp Ile Gly Gly Ser Ile Leu Ala Ser Leu Ser Thr Phe Gln						
	260		265		270	
Gln Met Gln Met Trp Ile Ala Lys Ala Glu Tyr Asp Glu Ser Gly Pro						
	275		280		285	
Ser Ile Val His Arg Lys Cys Phe						
	290		295			